METHOD AND APPARATUS FOR

ENCODING/DECODING INTERLACED VIDEO SIGNAL

BACKGROUND OF THE INVENTION

[01]   This application claims the priority of Korean Patent Application No. 2003-6541, filed on February 3, 2003, in the Korean Intellectual Property Office, the disclosure of which is incorporated herein in its entirety by reference.

1. Field of the Invention

[02]   The present invention relates to a system for encoding/decoding an interlaced video signal, and more particularly, to a video encoding/decoding method based on motion estimation and motion compensation of an interlaced video signal, and a video encoding/decoding apparatus.

2. Description of the Related Art

[03]   Typical MPEG-2 transcoders adaptively use frame motion estimation and field motion estimation to encode an interlaced video. Also, the H.264 recommendation, which is under standardization at the present, considers the encoding of an interlaced moving image.

[04]   FIG. 1 is a conceptual diagram of conventional motion estimation and motion compensation using two frames in an interlaced video. $F^t(n)$ and $F^b(n)$ denote a top field and a bottom field, respectively, of an n-th frame. It is

assumed that a current frame is an (n+1)th frame. For convenience, frames of the input video signal are shown on a time direction. In FIG. 1, a block to be motion-estimated, namely, a macroblock (MB), is composed of 8 vertically arranged pixels. A block to be motion-estimated of the current frame undergoes 5 motion estimations in a forward direction, namely, frame-to-frame motion estimation, top-to-top field motion estimation, top-to-bottom field motion estimation, bottom-to-top field motion estimation, and bottom-to-bottom field motion estimation. If a bi-directional motion estimation is required as in an MPEG-2 bi-directional picture, 5 forward motion estimations and 5 backward motion estimations are performed on the to be motion-estimated block of the current frame. Here, only forward motion estimation will be considered for the sake of convenience.

[05]   Referring to FIG. 1, in frame motion estimation (ME)/motion compensation (MC), an MB of a current frame F(n+1), indicated by a rectangular box, is matched with an MB of a reference frame F(n), indicated by a rectangular box, to find a frame motion vector MV_frame with a minimum Sum of Absolute Differences (SAD).

[06]   In top field ME/MC, a current top field $F^t(n+1)$ is matched with a reference top field $F^t(n)$ to find a motion vector MV_{t2t} with a minimum SAD_{t2t}. Also, the current top field $F^t(n+1)$ is matched with a reference bottom field $F^b(n)$ to find a motion vector MV_{t2b} with a minimum SAD_{t2b}.

[07]   In bottom field ME/MC, a current bottom field $F^b(n+1)$ is matched with the reference top field $F^t(n)$ to find a motion vector MV_{b2t} with a

minimum SAD_{b2t}. Also, the current bottom field $F^b(n+1)$ is matched with the reference bottom field $F^b(n)$ to find a motion vector MV_{b2b} with a minimum SAD_{b2b}.

[08]   A motion vector having SAD_{t2t} and a motion vector having SAD_{t2b} are compared, and a motion vector with a smaller SAD is determined to be a top field motion vector MV_{top_fld}. A motion vector having SAD_{b2t} and a motion vector having SAD_{b2b} are compared, and a motion vector having a smaller SAD is determined to be a bottom field motion vector MV_{bot_fld}. Hence, motion vectors to be used upon frame MC and field MC are all calculated by frame ME and field ME.

[09]   A SAD_{field} obtained from the top field motion vector MV_{top_fld} is compared with a SAD_{frame} obtained from the bottom field motion vector MV_{bot_fld}. If the SAD_{field} is smaller than the SAD_{frame}, field motion compensation occurs. If the SAD_{field} is greater than the SAD_{frame}, frame motion compensation occurs.

[10]   Such conventional frame ME/MC has the following problems. As shown in FIG. 2 (a), if the vertical component of a frame motion vector MV_{frame} (hereinafter, referred to as MV_{ver} is an even value, all of the pixels of a current macroblock have the same motion vector. Hence, frame motion compensation has no problems. However, if the MV_{ver} is an odd value as shown in FIG. 2 (b), frame motion compensation has a problem in that the motion vectors of the pixels corresponding to the top fields of the current macroblock are different from those of the pixels corresponding to the bottom

3

fields of the current macroblock. Thus, in the conventional frame ME/MC, the probability that the MV_{ver} is determined to be an even value increases, and unnecessary field motion compensation occurs due to inaccurate motion estimation and compensation. Therefore, unnecessary motion vector information may increase.

SUMMARY OF THE INVENTION

[11] The present invention provides a method of encoding/decoding an interlaced video signal, in which video motion estimation/compensation is performed in consideration of the actual locations of top field pixels and bottom field pixels that are received in an interlaced scanning way.

[12] The present invention provides a video encoding/decoding apparatus which performs the interlaced video encoding/decoding method according to the present invention.

[13] According to an aspect of the present invention, there is provided a video encoding/decoding method based on interlaced frame motion estimation and/or compensation. In the method, first, a macroblock and a search range are received, and a frame motion vector for each integer pixel is estimated using the received macroblock and search range. Then, if the vertical component of the estimated frame motion vector is an odd value, bottom field pixels in the received macroblock are matched with top field pixels in a reference frame that correspond to locations obtained by a scaled frame motion vector, whose vertical component has been scaled according to field-to-field distances. Also, top field pixels in the received macroblock are

4

matched with bottom field pixels in the reference frame that correspond to the original frame motion vector. On the other hand, if the vertical component of the estimated frame motion vector is an even value, the top or bottom field pixels in the received macroblock are matched with the top or bottom field pixels in the reference frame that correspond to the original frame motion vector.

[14] According to another aspect of the present invention, there is provided a method of encoding/decoding an interlaced video. In this method, first, a macroblock and a search range for image data are set. Then, it is determined whether the vertical component of a motion vector for each of integer pixels in the set macroblock is an even or odd value, and top and bottom field pixels in the macroblock are matched with field pixels in a reference frame that correspond to locations indicated by the motion vectors that are estimated different depending on the locations of pixels. Thereafter, if the motion vectors for the individual integer pixels of the macroblock have been completely estimated, the top/bottom field pixels in the macroblock are matched with half pixels in the reference frame that correspond to the motion vectors, wherein the matching is performed according to the vertical components of the motion vectors.

[15] According to still another aspect of the present invention, there is provided an apparatus for encoding an interlaced video. This apparatus includes a discrete cosine transform unit, a quantization unit, a dequantization unit, an inverse discrete cosine transform unit, a frame memory, and a motion

5

estimation/motion compensation unit. The discrete cosine transform unit performs a discrete cosine transform operation on individual macroblocks of incoming image data. The quantization unit quantizes the discrete cosine transformed image data. The dequantization unit dequantizes the quantized image data. The inverse discrete cosine transform unit performs inverse discrete cosine transform operation on the dequantized image data. The frame memory stores the inverse discrete cosine transformed image data on a frame-by-frame basis. The motion estimation/motion compensation unit determines whether the vertical component of a motion vector for each integer pixel in a macroblock is an even or odd value when the incoming image data of a current frame is compared with the image data of a previous frame stored in the frame memory. If the vertical component of the motion vector is an odd value, bottom field pixels are matched with top or bottom field pixels in the previous frame that correspond to motion vectors that are scaled depending on a distances between fields to be matched.

BRIEF DESCRIPTION OF THE DRAWINGS

[16] The above and other features and advantages of the present invention will become more apparent by describing in detail exemplary embodiments thereof with reference to the attached drawings in which:

[17] FIGS. 1 and 2 are conceptual diagrams of conventional motion estimation and motion compensation using two frames in an interlaced video;

[18] FIG. 3 is a block diagram of an interlaced video encoding apparatus according to the present invention;

6

[19]   FIG. 4 is a detailed flowchart for illustrating frame motion estimation (ME)/ motion compensation (MC) performed in the ME/MC unit of FIG. 3;

[20]   FIG. 5 is a detailed flowchart for illustrating frame ME when the vertical component of the motion vector of FIG. 4 is an odd value; and

[21]   FIG. 6 shows motion estimation and motion compensation using two frames in an interlaced video, according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[22]   Referring to FIG. 3, in an interlaced video encoding system according to the present invention, an incoming image corresponds to a group of pictures (GOP). A discrete cosine transform (DCT) unit 320 performs DCT on 8×8 blocks to obtain spatial redundancy from the incoming image and outputs a discrete cosine transformed (DCTed) image.

[23]   A quantization (Q) unit 330 quantizes the DCTed image. A dequantization unit 350 dequantizes the quantized image.

[24]   An inverse DCT (IDCT) unit 360 performs IDCT on the dequantized image and outputs an inverse DCTed (IDCTed) image. Frame memory (FM) 370 stores the IDCTed image on a frame-by-frame basis.

[25]   A motion estimation (ME)/motion compensation (MC) unit 380 estimates a motion vector (MV) for each macroblock and a Sum of Absolute Differences (SAD) by using image data of a current frame and image data of a previous frame stored in the frame memory 370, and performs motion compensation using the MVs.

[26]　A variable length coding (VLC) unit 340 removes statistical redundancy from the quantized image based on the MVs estimated by the ME/MC unit 380.

[27]　An interlaced video decoding apparatus restores a variable length coded (VLCed) image signal received from the interlaced video encoding apparatus to the original image signal by performing variable length decoding, dequantization, IDCT, and motion compensation.

[28]　FIG. 4 is a detailed flowchart for illustrating frame ME/MC performed in the ME/MC unit 380 of FIG. 3.　First, incoming image data is composed of macroblocks.　In step 410, a search range is predetermined to perform motion estimation on the macroblocks.

[29]　Next, a frame motion vector MV_{frame} for each integer pixel is estimated, in step 420.

[30]　In step 430, it is determined whether the vertical component of a frame motion vector (hereinafter, referred to as MV_{ver}) is an odd or even value.　If the MV_{ver} is an even value, conventional frame ME/MC occurs.

[31]　If the MV_{ver} is an odd value, motion vectors of pixels corresponding to the bottom and top fields in a current macroblock are calculated in different ways depending on the actual locations of the pixels, and the pixels in the current macroblock are matched with those in a reference frame according to the calculated motion vectors, in step 440.　In other words, SADs between the top field pixels in the current macroblock and the bottom field pixels in the reference frame are calculated by using the original MV_{ver} without change.

The bottom field pixels in the current macroblock are matched with top field pixels of the reference frame that are adjacent to the bottom field pixels, among the pixels with MV_{ver} properly scaled in consideration of the direction of an actual motion vector, obtaining SADs therebetween.

[32]   Thereafter, integer-pixel motion estimation is continuously performed on the next macroblock. In step 460, if there are no more macroblocks to be motion-estimated, it is considered that the integer-pixel motion estimation has been completed.

[33]   In step 470, the integer-pixel ME/MC is followed by ME/MC with respect to a half pixel (hereinafter, referred to as halfpel) or smaller pixels. For convenience, half-pixel motion estimation will be taken as an example hereinafter. In other words, if the MV_{ver} is an even value, all of the pixels in the next macroblock undergo general halfpel motion estimation. If the MV_{ver} is an odd value, the top field pixels in the next macroblock undergo halfpel ME/MC using bi-linear interpolation, and the bottom field pixels in the next macroblock are matched with pixels corresponding to the scaled MV_{ver}, and the matched pixels undergo halfpel ME/MC.

[34]   After the halfpel ME/MC has been completed, the integer-pixel MV_{ver} may not be distinguished whether it was an odd or even value. Accordingly, in step 480, when a frame MC mode has been selected for each macroblock, 1-bit information about whether the integer-pixel MV_{ver} is an odd or even value is produced. Thus, a decoder can decode image data with reference to the information about whether the MV_{ver} is an odd or even value.

[35]   Existing frame motion compensation, existing field motion

compensation, and frame motion compensation according to the present

invention can be adaptively used.

[36]   A decoder also can perform video motion estimation/compensation

according to the present invention.   In other words, the decoder performs

motion estimation/compensation on a video in consideration of the actual

locations of top field pixels and bottom field pixels, which are input according

to information about whether the MV_{ver} is an odd or even value, which is

received from an encoder.

[37]   FIG. 5 is a detailed flowchart for illustrating frame ME that occurs

when the MV_{ver} of FIG. 4 is an odd value.   For convenience, referring to FIG.

6, an incoming image signal is in a time direction.   A block to be motion-

estimated, namely, a macroblock (MB), is composed of vertically arranged 8

pixels. $F^t(n)$ and $F^b(n)$ denote a top field and a bottom field, respectively, of

an n-th frame. $F^t(n+1)$ and $F^b(n+1)$ denote a top field and a bottom field,

respectively, of an (n+1)th frame.   It is assumed that the (n+1)th frame is a

current frame.

[38]   First, in step 510, pixels that form a macroblock are input.

[39]   In step 520, it is determined whether the pixels belong to either bottom

or top fields.

[40]   In step 530, pixels corresponding to the top fields of the input

macroblock are matched with pixels corresponding to the bottom fields of a

reference frame, and an SAD between the former and latter pixels is obtained by using the original MV_{ver} without change.

[41]   In step 540, pixels corresponding to the bottom fields of the input macroblock are matched with pixels corresponding to the top fields of the reference frame, and an SAD between the former and latter pixels is obtained using the MV_{ver} that has been scaled. In other words, as shown in FIG. 6, the SAD between the pixels belonging to the bottom fields of the input macroblock and the pixels belonging to the top fields of the reference frame is obtained using a motion vector $a*MV_{ver}$, which is extended by a in consideration of the distances between matched fields. If the distance between $F^b(n)$ and $F^t(n+1)$ is d_{b2t}, and the distance between $F^t(n)$ and $F^b(n+1)$ is d_{t2b}, a is determined to be d_{b2t}/d_{t2b}. In FIG. 6, a location pointed by the motion vector $a*MV_{ver}$ is represented by x.

[42]   The top field pixels in the input macroblock are matched with the bottom field pixels in the reference frame that correspond to the motion vector MV_{ver}.

[43]   The pixels at the locations x shown in FIG. 6 can be either integer pixels or non-integer pixels. Accordingly, each of the locations x pointed by the motion vector $a*MV_{ver}$ is estimated using a top field pixel that is the closest to the location x. In other words, if the distance (d_u) between the pixel at the location x, P_x, and the integer pixel P_u right above the pixel P_x is smaller than or equal to the distance (d_d) between the pixel P_x and the integer pixel P_d right below the pixel P_x, the integer pixel P_u is selected as the pixel P_x. On the

other hand, if the distance (d_u) is greater than the distance (d_d), the integer pixel P_d is selected as the top field pixel to be matched with the bottom field pixel in the reference frame. Alternatively, if the distance (d_u) is smaller than the distance (d_d), the integer pixel P_u is selected as the top field pixel to be matched with the bottom field pixel in the reference frame. On the other hand, if the distance (d_u) is greater than or equal to the distance (d_d), the integer pixel P_d is selected as the pixel P_x.

[44]   Each of the locations x pointed by the motion vector $a*MV_{ver}$ can also be estimated using a bottom field pixel that is the closest to the location x.

[45]   While the present invention has been particularly shown and described with reference to exemplary embodiments thereof, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope of the present invention as defined by the following claims.

[46]   While the present invention has been particularly shown and described with reference to exemplary embodiments thereof, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope of the present invention as defined by the following claims. The invention can also be embodied as computer readable codes on a computer readable recording medium. The computer readable recording medium is any data storage device that can store data, which can be thereafter read by a computer system. Examples of the computer readable recording medium include read-only

12

memory (ROM), random-access memory (RAM), CD-ROMs, magnetic tapes, floppy disks, optical data storage devices, and so on. Also, the computer readable code can be transmitted via a carrier wave such as the Internet. The computer readable recording medium can also be distributed over a computer system network so that the computer readable code is stored and executed in a distributed fashion.

[47] As described above, video motion estimation/compensation according to the present invention is performed in consideration of the actual locations of the top field pixels and bottom field pixels that are received in an interlaced scanning way. Thus, the performance of motion compensation is improved, and the amount of motion vector information is reduced.